

Speaking Practice Support System for Japanese Learners of English and Investigation of Learning Strategies and Patterns through Learning Log Data

Yuichi Ono^{*} and Fu Suro[†]

Abstract

This paper constructs the system to support Japanese English learners to speak fluently and accurately with the use of Text-to-Speech (TTS) technology. In addition, this paper analyzes how the learners' learning behavior is changed according to the difficulty of the materials and investigates the relationship between their learning strategies and learning behaviors on the basis of learning log data. This paper concludes that the system affected their learning strategies according to the difficulty of the materials and their learning attitudes became more flexible and adaptive due to the use of the system.

Keywords: Speaking Practice, Learning Analytics, Learning Behavior, Learning Strategies

1 Introduction

This paper constructs a system which supports speaking practice for Japanese English learners of English as a Foreign Language (EFL), and investigates the relationship between learners' learning strategies and learning behavior patterns on the basis of the log data collected by the system. It is generally believed that productive skills like speaking in a foreign language are essential in the age of globalization. The effect is that recent curricula of English courses in higher education are focusing on presentation or speaking in public.

On the other hand, due to recent advancements of information technology, the accuracy and reliability of speech recognition and speech synthesis technology has been improving. Especially, educational applications of Text-to-Speech (TTS) technology is believed to help those who have speaking and listening disabilities as well as those who need listening and speaking practice [1][2].

Another recent trend of educational technology is the methodology of importing "big educational" data into analysis to predict learning outcomes and produce adaptation, personalization, and intervention to learning [3]. Since the system we constructed is able to store learners'

^{*} University of Tsukuba, Ibaraki, Japan

[†] Graduate School Student, University of Tsukuba, Ibaraki, Japan

learning log data on speaking practice, the analysis of the data can lead to the prediction of learners' learning patterns and, hopefully, their learning outcomes.

After reviewing backgrounds of this study, this paper describes the main concepts of the system construction [5], and lays out our research design which tries to investigate the relationship between learning patterns and learning log data. It will be concluded that the learners (i.e., participants of this study) will be rightly clustered in three groups in our clustering analysis. The clustered three groups are named as "Sentence-focused", "Word/Chunk Focused", and "Balanced". As to the relationship between learning behaviors and learning strategies, it will be demonstrated that some factors of learning strategies showed correlation with learning behaviors at the beginning of the course, but these correlation disappeared by the end of this project. This means that their initial learning patterns, which were triggered by their inherent learning strategies, were affected by the system-based speaking training.

2 Backgrounds

2.1 Needs for the System

It has been frequently requested from industries that the opportunities to practice speaking English more fluently and accurately should be provided with EFL learners so that they can listen to the model speech in preparation for academic or business presentation in English. If they had native speakers of English nearby when they need to practice, they could ask for help. With the supporting system, which provides the opportunities to practice sentences, chunks, and words they want to listen to, it is a good opportunity in which they can practice speaking at anytime, anywhere.

The model speech has been usually recorded by a native speaker in a tape or a CD. In this format, it is necessary that the model speech is pre-recorded. Furthermore, it is very difficult and troublesome to cut off some unfamiliar chunks or words to repeat during the practice with these devices. With the TTS system, it becomes possible to select the parts on the screen easily and have them speak soon after they insert the text.

2.2 Merits of TTS-Based System

TTS is one of the technologies considered to have a potential impact on education, which can give us the opportunity to read or listen to any text [2]. It can help learners to develop their listening or speaking skills [5][6].

TTS technology can be integrated into the classroom environment. In a view of pronunciation training, [2] indicates that the integration of TTS software into classrooms can help learners of English improve their pronunciation if the instructor can create time for practice sessions in the class. On the part of the instructors, TTS can be a tool for pronunciation training for instructors to prepare before lessons [7]. In addition, the use of TTS can decrease their anxiety about speaking in English in front of other students in sheltered practice sessions [2][5].

The discussions so far have outlined the possibility for improving EFL learners' oral speaking skills by constructing a TTS-based speaking practice system. With this system, students can practice accurate and fluent speaking with the help of TTS sounds. If they do

practice in sheltered speaking sessions or outside the classroom, they feel relaxed and less anxious during practice.

2.3 Quality of TTS

There are some studies validating the quality of TTS speech in Japanese EFL settings. In a study where TTS was used for both upper and lower proficiency groups in conducting listening tests, usefulness of using TTS in a listening test is shown, suggesting that participants in a lower group find it natural toward TTS, while those in upper group feel some unnaturalness toward it [8]. Another study analyzed the effects of TTS in the EFL classroom [9]. This study was conducted in a junior high school and a senior high school in Japan. The result was that there was a significant difference in speaking test between TTS users and non-TTS users. [10] built a web-based system using TTS for online courses, and further suggested its positive effects by evaluation, which showed 40% of students indicating that the voice of TTS is easy to listen to. Summarizing so far, the quality of TTS speech is generally fine, especially for learners with less proficiency, for speaking practice.

However, it is empirically clear that the quality of TTS speech depends on the type of software. Naturally expensive software will produce speech with more accuracy and fluency. In our system, Web Speech API [11] was employed for feasibility and into the system. According to our pilot survey of naturalness judgement test, with a combination of TTS speech and native speaker's speech, participants in our project answered a more degree of naturalness is observed in Word or Chunks, while they felt some degree of unnaturalness in Sentences. This is in accordance with previous studies.

2.4 Learning Analytics Approach

With the emergence of a huge amount of learning log data collected by new technologies, a field of research called Learning Analytics (LA) has started gaining attention recently. Learning analytics (LA) is “the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it takes place” [3]. Its goal is to gain a better understanding of how students learn and identify the setting in which they learn to improve educational outcomes and to gain insights into and explain educational phenomena [12].

In LA, such techniques as statistics, classification, visualization and sentiment statistics are commonly used for analysis [12]. In our study, clustering is employed on the basis of “Click” data of Word/Chunk/Sentence/Passage during the practice in order to produce grouping of different learning behavior types.

2.5 Learning Strategies and Learning Behaviors

When a learner works on a task of solving a question, various types of learner's factors are related to produce their own learning behavior. Learning strategy is believed to be one of the main factors which affect learning outcomes [13]. This is illustrated in Figure 1 below. Thus, the fundamental hypothesis of LA is that learning behaviors are a reflection of learner's unique learning strategies, learning beliefs, or other learner's factors. In our study

on speaking practice, our empirical assumptions are that there are two types of approach to work in speaking practice; (a) Top-Down Approach and (b) Bottom-Up Approach.

Top-Down learners generally start with larger units like Passages or Sentences, followed by some smaller units like Words or Chunks. This type of learners sometimes will pay little attention to difficult words or chunks in terms of accuracy. On the other hand, Bottom-Up learners start with Words or Chunks for precise pronunciation, then move to larger units.

Our tentative assumption is illustrated in Table 1 below:

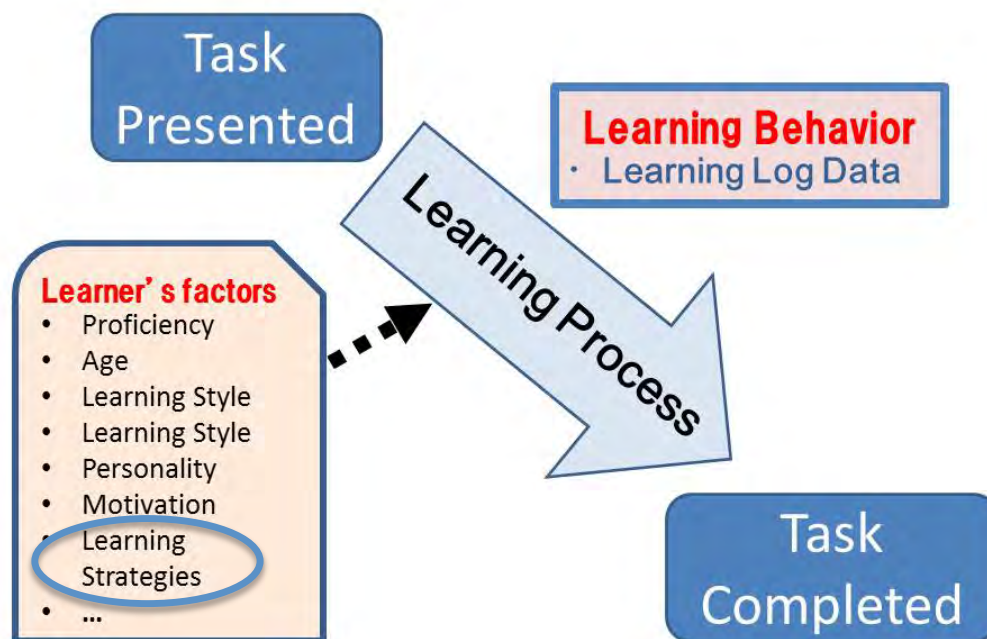


Figure 1: Scheme of Learning Process

Table 1: Initial Assumptions

| | Accuracy | Fluency |
|--------------------|----------|---------|
| Top-Down Learners | ?? | ◎ |
| Bottom-Up Learners | ◎ | ?? |

3 System Description

A rough image of our system design is shown in Figure 2 below.

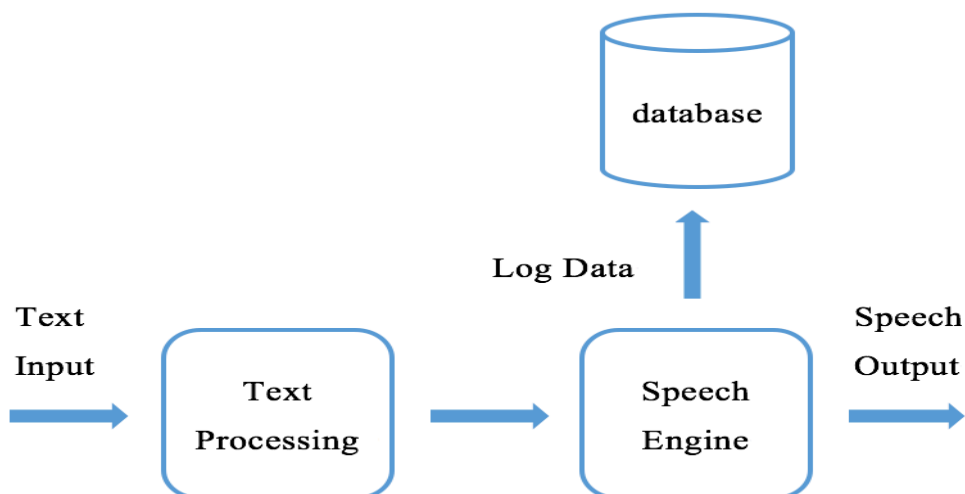


Figure 2: System Design

The system employs Web-Speech API for TTS, which involves Google Chrome extension. Data on the following will be stored for one “Click”.

- What to click
- When to click
- What mode to click (Word/Chunk/Sentence/Passage)

After each lesson, students are request to check (click) Unknown Words (UWs) and Difficult to Pronounce Words (DPWs) for reflection of the activity. The interface screenshot is given in Figure 3 below:

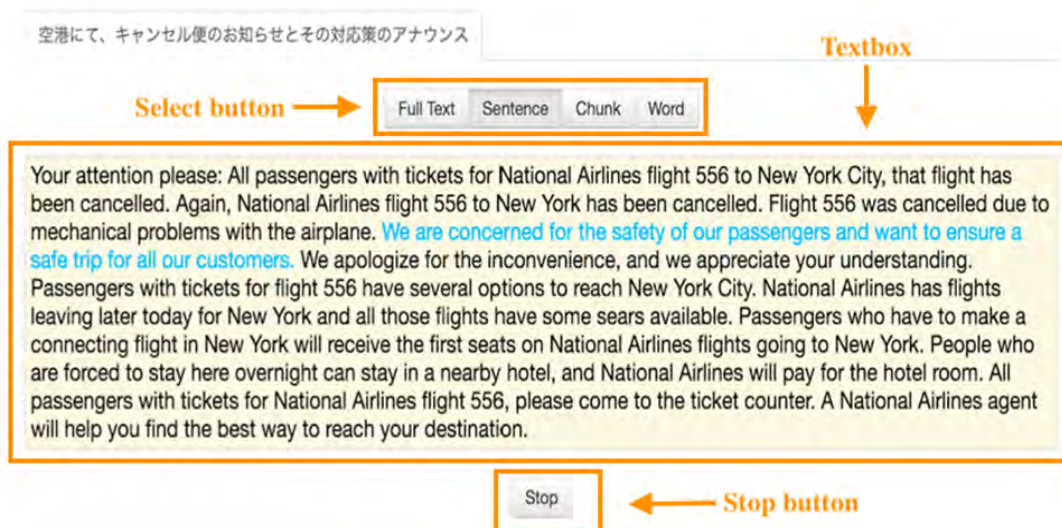


Figure 3: Interface

We have a “Select” button to choose the mode; Full Text, Sentence, Chunk or Word. This picture shows that the student chooses “Sentence Mode” to practice. And when they click part of the sentence, the sentence is automatically selected and the speech appears. When they finish the practice, they click the “Stop” button.

As to the visualization, our system will automatically visualize data using Google Charts API, which can help us to create charts in a simple way. The following figures show when and how they practiced speaking. Figures 4 and 5 are examples of Top-Down and Bottom-Up Learners, respectively.

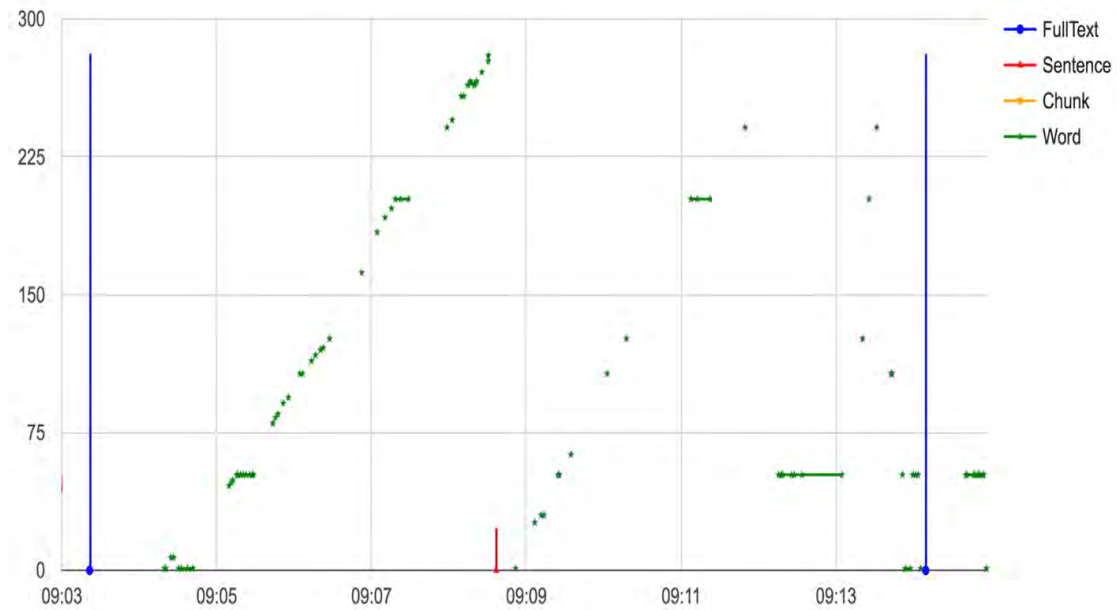


Figure 3: Visualization (Top-Down Learner's Behavior)

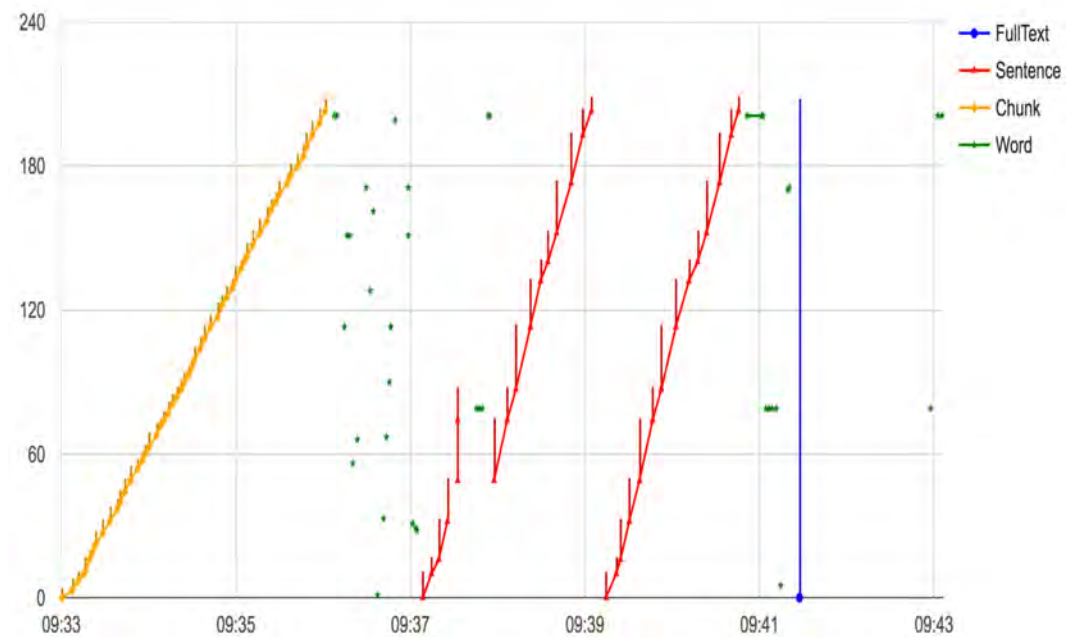


Figure 4: Visualization (Bottom-Up Learner's Behavior)

We can see from Figure 3 that this learner starts with Full-text listening, then moves to Words maybe checking pronunciations. On the other hand, the Bottom-Up learner in Figure 4 starts with Chunks, then some Words, then Sentence, and lastly, ends with Full-text. This is a direct reflection to learners, which has a possibility for deep reflections on their learning behavior and outcomes. In our previous studies [4], it was shown that almost all of the students showed improvement in the speaking test and that, according to our questionnaire research, they were satisfied with the speaking practice study in terms of interface and the quality of the feedback.

4 Studies

This chapter deals with the following two research questions.

RQ1: How are the learners divided as a result of clustering analysis and how are they changed?

RQ2: Is there any correlation between learner's strategies and learning log data?

4.1 Procedure and Instrument

A total of 82 first-year students in a Japanese national university participated in this study. On the basis of result of placement test, 43 students are assigned to the Upper-level class, and 39 students to the Lower class. We constructed seven materials (Lesson 1, 2, ..., 7) for seven weeks. In order to see how they change their learning patterns, we mixed both easy and difficult materials for practice. In each class, they are told to record their speech before and after the practice with the system.

As to RQ1, we conducted Cluster analysis involving Ward method with variances of (a) frequency of jump in Sentence mode, (b) frequency of return in Sentence mode, (c) frequency of jump in Chunk mode, and (d) frequency of return in Chunk mode. Observing dendrograms for each lessons carefully, we determined three groups for each lesson.

As to RQ2, a questionnaire research is conducted to ask learners' learning strategies at the start and end of the project. We employed Strategies Inventory for Language Learning (SILL) consisting of six categories with 50 question items [13]. The categories are (a) Cognitive Strategies, (b) Metacognitive Strategies, (c) Memory-related Strategies, (d) Compensatory Strategies, (e) Affective Strategies, and (f) Social Strategies. We conducted Spearman's Correlation analysis to observe a correlation between learning strategies and learning behavior.

4.2 Results

The result of clustering is given in Figure 5, where No's 2, 4, 5, and 6 are easier materials as to the readability figures. According to the results of our cluster analysis, we divided learners into the following three types; *Sentence Centered* (i.e., Top Down), *Chunk/Word Focused* (i.e., Bottom-Up), and *Balanced* (i.e., Mixed). You can see that learners seem to change their learning behaviors according to the contents of the materials. Looks like no difference between Upper Class and Lower Class. Thus, the behaviors are less relevant to their proficiencies. Observing the difficult three materials (that is, No.1, No.3, and No.7), we can see that there are a large number of learners belonging to Balanced or Word/Chunk groups. These things seem to suggest that learners pay more attention to Words/Chunks when materials are difficult for accuracy

rather than fluency. This is consistent with the results of our earlier pilot study [4].

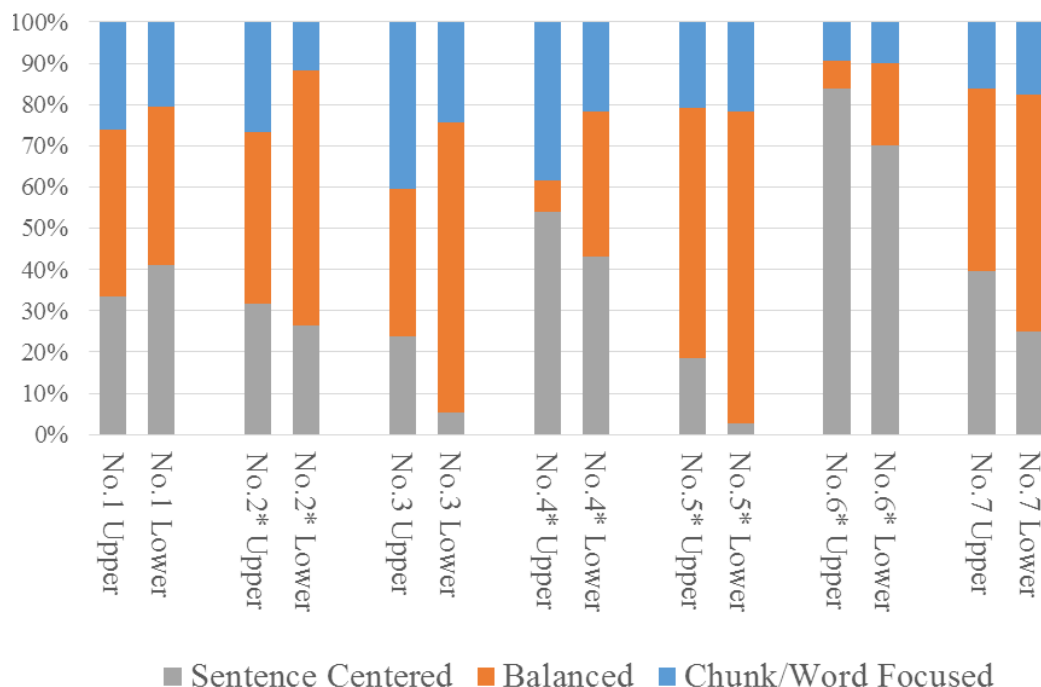


Figure 5: Percentages of Each group

As to RQ2, the result says that there are only two categories of SILL which showed significant correlation with learning behavior before the project. However, these correlations were not observed after the project. This is shown in Table 2 below. This result seems to suggest that the initial learning strategies are affected by the system-based speaking practice. This is an important implication for our perspective, because this might mean that their learning strategies became more flexible according to the quality of materials, and this is a sign of successful learners who can speak accurately and fluently with wider variety of strategy use [13].

Table 2: Correlation Analysis

| | Start | | End | |
|-------------------------------------|----------|----------|----------|-----------|
| | <i>r</i> | <i>p</i> | <i>r</i> | <i>p</i> |
| Social Strategies – Full-text | 0.259 | 0.022 | 0.008 | <i>ns</i> |
| Social Strategies – Chunk | -0.238 | 0.036 | -0.108 | <i>ns</i> |
| Metacognition Strategies – Sentence | 0.253 | 0.025 | 0.186 | <i>ns</i> |

5 Concluding Remarks

This paper described the system which supports Japanese EFL learner’s speaking practice with the use of TTS technology. On the basis of the learning log data collected by the system,

we analyzed learners' learning behavior patterns. The analysis showed that there are three groups constructed by our analysis and the learners changed their behavior patterns according to the difficulty of materials. As to the relationship between learning strategies and learning behavior, there are some correlations between them at the beginning of the course. However, this relationship was not observed at the end of the course. This might be related to the situation that their learning strategy use became flexible and more adaptive so that they can learn to speak accurately and fluently.

Acknowledgement

This study was partially supported by Grant-in-Aid for Scientific Research No. 26370656 from the Japan Society for the Promotion of Science (JSPS).

References

- [1] D. González, "Text-to-speech applications used in EFL contexts to enhance pronunciation," *TESL-EJ*, 11(2), 2007, pp.1-11.
- [2] F. Kilickaya, "Text-To-Speech Technology': What Does It Offer To Foreign Language Learners?" *CALL-EJ Online*, 7(2), 2007.
- [3] R. Ferguson, "Learning analytics: drivers, developments and challenges," *International Journal of Technology Enhanced Learning*, 4(5-6), 2012, pp. 304-317.
- [4] F. Suro and Y. Ono, "Japanese EFL Learners' Uses of Text-to-Speech Technology and Their Learning Behaviors: A Pilot Study," 2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), Kumamoto, 2016, pp. 296-301.
- [5] K. Aoki, "Effective Use of Text-to-Speech Technology for Japanese Learners of English," *NTUT Education of Disabilities*, 7, 2009, pp.1-10.
- [6] Y. Zhao, "Recent developments in technology and language learning: A literature review and meta-analysis," *CALICO journal*, 2003, pp.7-27.
- [7] J. Azuma, "Impact of TTS technology on foreign language teaching: New horizons of multimedia teaching material development," *Ryutsu Kagaku Daigaku Kyoiku Koudoka Suishin Center Kiyuu*, 6, 2010, pp.1-11.
- [8] A. Hirai and T. Oki, "Comprehensibility and Naturalness of Text-To-Speech Synthetic Materials for EFL Listeners," *JACET Journal*, (53), 2011, pp.1-17.
- [9] S. Fujishiro and I. Miyaji, "A Study on the Utilization of the English Text-to-Speech Software in Blended Instruction and Its Effect," *JSET*, 32(4), 2009, pp.395-404.
- [10] H. Kashiwagi, M. Kang, and K. Ohtsuki, "Current Status and Future Prospects of Application of Synthetic Speech in Foreign Language Learning," *Kokusai Communication Center Ronsyuu*, 5, 2008, pp.10-19.

- [11] Web Speech API: <https://developer.mozilla.org/ja/docs/Web/API/SpeechSynthesis>
- [12] C. Romero and S. Ventura, "Data mining in education," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 2013, pp.12-27.
- [13] R. Oxford, "Language learning strategies: What every teacher should know," Boston: Heinle, 1990.